

EBPM の実現に向けた都市の特徴についての分析

(株)国土開発センター	正会員	○中宮 雅貴	(株)国土開発センター	非会員	新家 哲平
(株)国土開発センター	非会員	稲田 裕介	(株)国土開発センター	非会員	新田川貴之
(株)国土開発センター	非会員	森川 大輔	金沢工業大学	非会員	武市 祥司

1. 研究の目的

近年、ICT の進展により、ビッグデータおよびオープンデータの活用や、AI の実用化に向けた取組が進んでいる。国や地方公共団体等においてもデジタル・ガバメントの実現に向けた取組やオープンデータを活用した分析が推進されており、データ活用による経済や地域の活性化等が期待されている。

本研究は、地方公共団体等において、EBPM（根拠に基づく政策立案）を推進する際の検討手法としての活用を目的とする。本稿では、「都市の特性を表す様々な統計が整理されたオープンデータである都市モニタリングシート（国土交通省公表）を活用した多変量解析による都市の特徴についての分析」について述べる。

2. 研究概要

2-1 活用データ

全国 1,719 都市（2017 年度の市区町村数、東京都区部は 1 都市とする）の基礎情報（人口・面積等）・交通・防災・産業・経済・財政・生活の利便性の指標・健康・福祉の指標などの 351 個の指標が記載されたオープンデータである都市モニタリングシートを用いる。

2-2 分析の流れ

①データ前処理

351 個の指標から、指標の性質や欠損値の少なさ等を判断基準に 86 個を選別した上で、欠損値補完および標準化（z-score normalization）を行った。

②データの集約

多くある指標のうち、似た傾向のものを統合して集約するため、指標に影響を与える潜在因子を見出す因子分析を行い、86 指標から 17 個の因子を見出し、うち因子の性質が説明可能な 9 個を選別した。（図 1）

①生活利便性	④農業の振興（活力）	⑦土砂災害危険度
②都市の規模	⑤事業費（事業規模）	⑧富裕度合い
③都市の資産	⑥人口増減と財政負担	⑨地価変化

図 1 因子の一覧

③都市の分類

都市の特徴を捉えやすくし、後工程で行う重回帰分析の精度を高めるため、②で選別した因子の値を基準に、同傾向を示す都市群を分類するクラスター分析（k-means 法）を行った。

分析の結果、全国 1,719 都市が 18 個の都市群に分類され、因子の値から、都市群に名前付けを行った。「生活利便性」や「都市の規模」、「富裕度合い」等の因子が主要ファクターとなっている都市群の他、これ以外の因子がファクターとなっている都市群も見られた。一方で、因子値から名前付けし難い都市群も見られた。

④相関分析・モデル検証（重回帰分析）

③で分類された都市のデータを用いて、重回帰分析により指標と因子との関係性を明らかにする。詳細は第 3 章で詳述する。

⑤都市の特徴の分析

④までの分析結果から、特定の都市に着目し、その特徴を見出す。具体的には、重回帰分析により得られた都市のある指標に対する現況値と予測値の差分に着目し、その差分に及ぼす要因に関して分析・考察する。

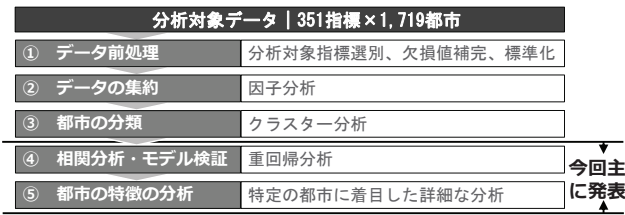


図 2 分析の流れ

3. 分析結果

3-1 相関分析（重回帰分析）

都市モニタリングシートの 86 種類の指標データそれぞれに対して、9 個の因子で説明する重回帰式を求めた。重回帰式を求める際には、以下の点についての配慮を行っている。

- ・指標と因子の組合せの中には、関係性が低いものも存在するため、ステップワイズ法により説明変数（因子）の選択を行う。

3-2 回帰モデル検証（F検定）

回帰式の有意性をF検定により判定した（図3）. 一部 $\alpha = 0.05$ を満たさない回帰式も存在したが、9割以上の回帰式は1%有意であった.

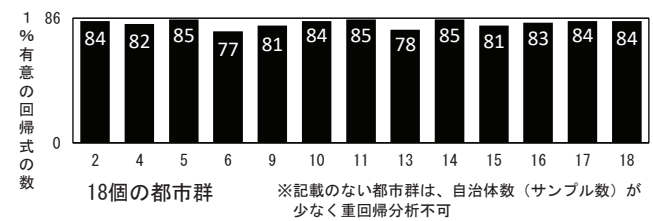


図3 F 検定結果（1%有意のモデル数）

3-3 都市の特徴的分析

重回帰分析で求めた回帰式の説明変数に、調べたい自治体（市町村）が持つ因子の値を投入することで、その自治体の指標の予測値が算出できる. この指標の現況値と予測値を比較することにより、その自治体が、同じクラスターに属する類似自治体の中で、どの位置づけにあるかを把握することが可能と考える.

図4は、北陸地方のN市（クラスター分析結果から「地方の衛星都市・利便性の高い都市」クラスターに分類される）を対象に、現況値と予測値の差をプロットしたサンプルデータである.

図4から、以下のようなことが読み取れる.

- ・高専が所在する市であるため、「No.1：高専数」が突出している.
- ・「No.6：自動車分担率」が高い一方で「No.84：鉄道分担率」が低い.
- ・「No.2：日常生活徒歩圏」や「No.5：駅バス停徒歩圏カバー率」、「No.7：商業施設徒歩圏カバー率」が比較的高い一方で、「No.83：福祉施設徒歩圏カバー率」、「No.86：保育所徒歩圏」等は低い.

これらのことからN市は、「地方の衛星都市・利便性の高い都市」都市群の中でも、項目によっては高い住民利便性を有する自治体であるが、福祉施設や保育所が少ない、車への高依存などの現状がうかがわれる.

4. 今後の課題と展望

本研究では、特徴を洗い出したい任意の都市が存在する場合に、オープンデータを活用して、因子分析・クラスター分析・重回帰分析を組み合わせることで、

特徴を洗い出す、1つの手法を提示した.

今回の重回帰分析では、特定の都市に着目した分析を行ったが、分析結果の有用性を検証するために、今後は様々な都市でケーススタディを行いたい.

本稿では、重回帰分析において、説明変数として指標を集約した因子を設定したが、説明変数と目的変数の関係性（目的変数に強い影響を及ぼす変数）を分析する場合においては、指標を説明変数に用いた方がより実態に即した分析が可能と考える. このため、説明変数・目的変数の両者に指標を用い、分析上有意な関係性を明らかにし、指標の推計シミュレーションを行ってみたい.（例えば、公共交通の利便性（説明変数）を高めることで、人口（目的変数）が増加するなど）

今後の展望としては、推計シミュレーションを進め、その結果を地方公共団体等の政策を検討する際のエビデンスとして活用することを目指したい. また、将来的には、都市モニタリングシート以外の統計データも分析対象に加え、より実態に即した分析を行いたい.



図4 差分プロット（北陸地方N市の例）

キーワード EBPM, ビッグデータ, オープンデータ, クラスター分析, 多変量解析, 重回帰分析

〒924-0838 石川県白山市八東穂 3-7（石川ソフトリサーチパーク内） TEL:076-274-8818, FAX:076-274-8421